Sequence Listing

<110> LUKACSOVICH, Tamas
ASZTALOS, Zoltan
YAMAMOTO, Daisuke
AWANO, Wakae

<120> A Vector for Gene Trap, and A Method for Gene Trapping by Using The Vector

<130> 2000-1561A/LC/00653

<140> 09/700,843
<141> 2001-02-07

<150> PCT/JP99/02683
<151> 1999-05-21

<150> JP 10-141952
<151> 1998-05-22

<160> 7

<170> PatentIn Ver. 2.0

<210> 1
<211> 11206
<212> DNA
<213> Artificial sequence

<220>
<221> 3'P sequence
<222> (1)..(237)

<220>
<221> synthetic splicing acceptor site and stop/start sequence
<222> (238)..(274)

<220>
<221> Gal4 gene (coding region and 3'UTR)
<222> (275)..(3164)

<220>
<221> hsp70 terminator
<222> (3165)..(3426)

<220>
<221> synthetic junction sequence
<222> 3427-3457

<220>
<221> heat shock promoter directed neomycine resistance gene on complementer strand
<222> (3458)..(4907)

<220>
<221> mini-white gene

<222> (4908)..(8275)

<220>
<221> synthetic splicing donor site
<222> (8276)..(8299)

<220>
<221> 5'P sequence
<222> (8300)..(8446)

<220>
<221> bacterial part of pCasper3 shuttle vector including complete pUC8 sequence
<222> (8447)..(11206)

<220>
<221> synthetic DNA
<222> (238)..(274)
(3427)..(3457)
(4908)..(4914)(8276)..(8299)
<223> Description of Artificial Sequence: Synthetic DNA

<400> 1
catgatgaaa taacataagg tggtcccgtc ggcaagagac atccacttaa cgtatgcttg 60
caataagtgc gagtgaaagg aatagtattc tgagtgtcgt attgagtctg agtgagacag 120
cgatatgatt gttgattaac ccttagcatg tccgtggggt ttgaattaac tcataatatt 180
aattagacga aattattttt aaagttttat ttttaataat ttgcgagtac gcaaagctct 240
ttctcttaca ggtcgaattg atgtgatgga tccaatgaag ctactgtctt ctatcgaaca 300
agcatgcgat atttgccgac ttaaaaagct caagtgctcc aaagaaaaac cgaagtgcgc 360
caagtgtctg aagaacaact gggagtgtcg ctactctccc aaaaccaaaa ggtctccgct 420
gactagggca catctgacag aagtggaatc aaggctagaa agactggaac agctatttct 480
actgattttt cctcgagaag accttgacat gattttgaaa atggattctt tacaggatat 540
aaaagcattg ttaacaggat tatttgtaca agataatgtg aataaagatg ccgtcacaga 600
tagattggct tcagtggaga ctgatatgcc tctaacattg agacagcata gaataagtgc 660
gacatcatca tcggaagaga gtagtaacaa aggtcaaaga cagttgactg tatcgattga 720
ctcggcagct catcatgata actccacaat tccgttggat tttatgccca gggatgctct 780
tcatggattt gattggtctg aagaggatga catgtcggat ggcttgccct tcctgaaaac 840
ggaccccaac aataatgggt tctttggcga cggttctctc ttatgtattc ttcgatctat 900
tggctttaaa ccggaaaatt acacgaactc taacgttaac aggctcccga ccatgattac 960
ggatagatac acgttggctt ctagatccac aacatcccgt ttacttcaaa gttatctcaa 1020
taattttcac ccctactgcc ctatcgtgca ctcaccgacg ctaatgatgt tgtataataa 1080
ccagattgaa atcgcgtcga aggatcaatg gcaaatcctt tttaactgca tattagccat 1140
tggagcctgg tgtatagagg gggaatctac tgatatagat gtttttttact atcaaaatgc 1200
taaatctcat ttgacgagca aggtcttcga gtcaggttcc ataattttgg tgacagccct 1260
acatcttctg tcgcgatata cacagtggag gcagaaaaca aatactagct ataattttca 1320
cagcttttcc ataagaatgg ccatatcatt gggcttgaat agggacctcc cctcgtcctt 1380
cagtgatagc agcattctgg aacaaagacg ccgaatttgg tggtctgtct actcttggga 1440
gatccaattg tccctgcttt atggtcgatc catccagctt tctcagaata caatctcctt 1500
cccttcttct gtcgacgatg tgcagcgtac cacaacaggt cccaccatat atcatggcat 1560
cattgaaaca gcaaggctct tacaagtttt cacaaaaatc tatgaactag acaaaacagt 1620
aactgcagaa aaaagtccta tatgtgcaaa aaaatgcttg atgatttgta atgagattga 1680
ggaggtttcg agacaggcac caaagttttt acaaatggat atttccacca ccgctctaac 1740
caatttgttg aaggaacacc cttggctatc ctttacaaga ttcgaactga agtggaaaca 1800
gttgtctctt atcatttatg tattaagaga ttttttcact aattttaccc agaaaaagtc 1860
acaactagaa caggatcaaa atgatcatca aagttatgaa gttaaacgat gctccatcat 1920
gttaagcgat gcagcacaaa gaactgttat gtctgtaagt agctatatgg acaatcataa 1980
tgtcacccca tattttgcct ggaattgttc ttattacttg ttcaatgcag tcctagtacc 2040

cataaagact ctactctcaa actcaaaatc gaatgctgag aataacgaga ccgcacaatt 2100
attacaacaa attaacactg ttctgatgct attaaaaaaa ctggccactt ttaaaatcca 2160
gacttgtgaa aaatacattc aagtactgga agaggtatgt gcgccgtttc tgttatcaca 2220
gtgtgcaatc ccattaccgc atatcagtta taacaatagt aatggtagcg ccattaaaaa 2280
tattgtcggt tctgcaacta tcgcccaata ccctactctt ccggaggaaa atgtcaacaa 2340
tatcagtgtt aaatatgttt ctcctggctc agtagggcct tcacctgtgc cattgaaatc 2400
aggagcaagt ttcagtgatc tagtcaagct gttatctaac cgtccaccct ctcgtaactc 2460
tccagtgaca ataccaagaa gcacaccttc gcatcgctca gtcacgcctt ttctagggca 2520
acagcaacag ctgcaatcat tagtgccact gaccccgtct gctttgtttg gtggcgccaa 2580
ttttaatcaa agtgggaata ttgctgatag ctcattgtcc ttcactttca ctaacagtag 2640
caacggtccg aacctcataa caactcaaac aaattctcaa gcgctttcac aaccaattgc 2700
ctcctctaac gttcatgata acttcatgaa taatgaaatc acggctagta aaattgatga 2760
tggtaataat tcaaaaccac tgtcacctgg ttggacggac caaactgcgt ataacgcgtt 2820
tggaatcact acagggatgt ttaataccac tacaatggat gatgtatata actatctatt 2880
cgatgatgaa gataccccac caaacccaaa aaaagagtaa aatgaatcgt agatactgaa 2940
aaaccccgca agttcacttc aactgtgcat cgtgcaccat ctcaatttct ttcatttata 3000
catcgttttg ccttctttta tgtaactata ctcctctaag tttcaatctt ggccatgtaa 3060
cctctgatct atagaatttt ttaaatgact agaattaatg cccatctttt ttttggacct 3120
aaattcttca tgaaaatata ttacgagggc ttattcagaa gcttatcgat accgtcgact 3180
aaagccaaat agaaattatt cagttctggc ttaagttttt aaaagtgata ttatttattt 3240
ggttgtaacc aaccaaaaga atgtaaataa ctaatacata attatgttag ttttaagtta 3300
gcaacaaatt gattttagct atattagcta cttggttaat aaatagaata tatttattta 3360
aagataattc gtttttattg tcagggagtg agtttgctta aaaactcgtt tagatccact 3420
agaaggaccg cggctcctcg accggatcga aaggagggcg aagaactcca gcatgagatc 3480
cccgcgctgg aggatcatcc agccggcgtc ccggaaaacg attccgaagc ccaacctttc 3540
atagaaggcg gcggtggaat cgaaatctcg tgatggcagg ttgggcgtcg cttggtcggt 3600
catttcgaac cccagagtcc cgctcagaag aactcgtcaa gaaggcgata gaaggcgatg 3660
cgctgcgaat cgggagcggc gataccgtaa agcacgagga agcggtcagc ccattcgccg 3720
ccaagctctt cagcaatatc acgggtagcc aacgctatgt cctgatagcg gtccgccaca 3780
cccagccggc cacagtcgat gaatccagaa aagcggccat tttccaccat gatattcggc 3840
aagcaggcat cgccatgggt cacgacgaga tcctcgccgt cgggcatgcg cgccttgagc 3900
ctggcgaaca gttcggctgg cgcgagcccc tgatgctctt cgtccagatc atcctgatcg 3960
acaagaccgg cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggtcg 4020
aatgggcagg tagccggatc aagcgtatgc agccgccgca ttgcatcagc catgatggat 4080
actttctcgg caggagcaag gtgagatgac aggagatcct gccccggcac ttcgcccaat 4140
agcagccagt cccttcccgc ttcagtgaca acgtcgagca cagctgcgca aggaacgccc 4200
gtcgtggcca gccacgatag ccgcgctgcc tcgtcctgca gttcattcag ggcaccggac 4260
aggtcggtct tgacaaaaag aaccgggcgc ccctgcgctg acagccggaa cacggcggca 4320
tcagagcagc cgattgtctg ttgtgcccag tcatagccga atagcctctc cacccaagcg 4380
gccggagaac ctgcgtgcaa tccatcttgt tcaatcatgc gaaacgatcc tcatcctgtc 4440
tcttgatcag atcccctatt cagagttctc ttcttgtatt caataattac ttcttggcag 4500
atttcagtag ttgcagttga tttacttggt tgctggttac ttttaattga ttcactttaa 4560
cttgcacttt actgcagatt gtttagcttg ttcagctgcg cttgtttatt tgcttagctt 4620
tcgcttagcg acgtgttcac tttgcttgtt tgaattgaat tgtcgctccg tagacgaagc 4680
gcctctattt atactccggc gctcttttcg cgaacattcg aggcgcgctc tctcgaacca 4740
acgagagcag tatgccgttt actgtgtgac agagtgagag agcattagtg cagagaggga 4800
gagacccaaa aagaaaagag agaataacga ataacggcca gagaaatttc tcgagttttc 4860
tttctgccaa acaaatgacc taccacaata accagtttgt tttgggatct agtccctaat 4920
tctagtatgt atgtaagtta ataaaaccct tttttggaga atgtagattt aaaaaaacat 4980
attttttttt tatttttttac tgcactggac atcattgaac ttatctgatc agttttaaat 5040
ttacttcgat ccaagggtat ttgaagtacc aggttctttc gattacctct cactcaaaat 5100
gacattccac tcaaagtcag cgctgtttgc ctccttctct gtccacagaa atatcgccgt 5160
ctctttcgcc gctgcgtccg ctatctcttt cgccaccgtt tgtagcgtta cctagcgtca 5220
atgtccgcct tcagttgcac tttgtcagcg gtttcgtgac gaagctccaa gcggtttacg 5280
ccatcaatta aacacaaagt gctgtgccaa aactcctctc gcttcttatt tttgtttgtt 5340
ttttgagtga ttggggtggt gattggtttt gggtgggtaa gcaggggaaa gtgtgaaaaa 5400
tcccggcaat gggccaagag gatcaggagc tattaattcg cggaggcagc aaacacccat 5460

ctgccgagca tctgaacaat gtgagtagta catgtgcata catcttaagt tcacttgatc 5520
tataggaact gcgattgcaa catcaaattg tctgcggcgt gagaactgcg acccacaaaa 5580
atcccaaacc gcaatcgcac aaacaaatag tgacacgaaa cagattattc tggtagctgt 5640
gctcgctata taagacaatt tttaagatca tatcatgatc aagacatcta aaggcattca 5700
ttttcgacta cattcttttt tacaaaaaat ataacaacca gatattttaa gctgatccta 5760
gatgcacaaa aaataaataa aagtataaac ctacttcgta ggatacttcg ttttgttcgg 5820
ggttagatga gcataacgct tgtagttgat atttgagatc ccctatcatt gcagggtgac 5880
agcggacgct tcgcagagct gcattaacca gggcttcggg caggccaaaa actacggcac 5940
gctcctgcca cccagtccgc cggaggactc cggttcaggg agcggccaac tagccgagaa 6000
cctcacctat gcctggcaca atatggacat ctttggggcg gtcaatcagc cgggctccgg 6060
atggcggcag ctggtcaacc ggacacgcgg actattctgc aacgagcgac acataccggc 6120
gcccaggaaa catttgctca agaacggtga gtttctattc gcagtcggct gatctgtgtg 6180
aaatcttaat aaagggtcca attaccaatt tgaaactcag tttgcggcgt ggcctatccg 6240
ggcgaacttt tggccgtgat gggcagttcc ggtgccggaa agacgaccct gctgaatgcc 6300
cttgcctttc gatcgccgca gggcatccaa gtatcgccat ccgggatgcg actgctcaat 6360
ggccaacctg tggacgccaa ggagatgcag gccaggtgcg cctatgtcca gcaggatgac 6420
ctctttatcg gctccctaac ggccagggaa cacctgattt tccaggccat ggtgcggatg 6480
ccacgacatc tgacctatcg gcagcgagtg gcccgcgtgg atcaggtgat ccaggagctt 6540
tcgctcagca aatgtcagca cacgatcatc ggtgtgcccg gcagggtgaa aggtctgtcc 6600
ggcggagaaa ggaagcgtct ggcattcgcc tccgaggcac taaccgatcc gccgcttctg 6660
atctgcgatg agcccacctc cggactggac tcatttaccg cccacagcgt cgtccaggtg 6720
ctgaagaagc tgtcgcagaa gggcaagacc gtcatcctga ccattcatca gccgtcttcc 6780
gagctgtttg agctctttga caagatcctt ctgatggccg agggcagggt agctttcttg 6840
ggcactccca gcgaagccgt cgacttcttt tcctagtgag ttcgatgtgt ttattaaggg 6900
tatctagcat tacattacat ctcaactcct atccagcgtg ggtgcccagt gtcctaccaa 6960
ctacaatccg gcggactttt acgtacaggt gttggccgtt gtgcccggac gggagatcga 7020
gtcccgtgat cggatcgcca agatatgcga caattttgct attagcaaag tagcccggga 7080
tatggagcag ttgttggcca ccaaaaattt ggagaagcca ctggagcagc cggagaatgg 7140
gtacacctac aaggccacct ggttcatgca gttccgggcg gtcctgtggc gatcctggct 7200
gtcggtgctc aaggaaccac tcctcgtaaa agtgcgactt attcagacaa cggtgagtgg 7260
ttccagtgga aacaaatgat ataacgctta caattcttgg aaacaaattc gctagatttt 7320
agttagaatt gcctgattcc acacccttct tagtttttt caatgagatg tatagtttat 7380
agttttgcag aaaataaata aatttcattt aactcgcgaa catgttgaag atatgaatat 7440
taatgagatg cgagtaacat tttaatttgc agatggttgc catcttgatt ggcctcatct 7500
ttttgggcca acaactcacg caagtgggcg tgatgaatat caacggagcc atcttcctct 7560
tcctgaccaa catgaccttt caaaacgtct ttgccacgat aaatgtaagt cttgtttaga 7620
atacatttgc atattaataa tttactaact ttctaatgaa tcgattcgat ttaggtgttc 7680
acctcagagc tgccagtttt tatgagggag gcccgaagtc gactttatcg ctgtgacaca 7740
tactttctgg gcaaaacgat tgccgaatta ccgctttttc tcacagtgcc actggtcttc 7800
acggcgattg cctatccgat gatcggactg cgggccggag tgctgcactt cttcaactgc 7860
ctggcgctgg tcactctggt ggccaatgtg tcaacgtcct tcggatatct aatatcctgc 7920
gccagctcct cgacctcgat ggcgctgtct gtgggtccgc cggttatcat accattcctg 7980
ctctttggcg gcttcttctt gaactcgggc tcggtgccag tatacctcaa atggttgtcg 8040
tacctctcat ggttccgtta cgccaacgag ggtctgctga ttaaccaatg ggcggacgtg 8100
gagccgggcg aaattagctg cacatcgtcg aacaccacgt gccccagttc gggcaaggtc 8160
atcctggaga cgcttaactt ctccgccgcc gatctgccgc tggactacgt gggtctggcc 8220
attctcatcg tgagcttccg ggtgctcgca tatctggctc taagacttcg ggcccgacgc 8280
aaggagtaga aggtaagtag cggccgcacg taagggttaa tgttttcaaa aaaaaattcg 8340
tccgcacaca acctttcctc tcaacaagca aacgtgcact gaatttaagt gtatacttcg 8400
gtaagcttcg gctatcgacg ggaccacctt atgttatttc atcatgggcc agacccacgt 8460
agtccagcgg cagatcggcg gcggagaagt taagcgtctc caggatgacc ttgcccgaac 8520
tggggcacgt ggtgttcgac gatgtgcagc taatttcgcc cggctccacg tccgcccatt 8580
ggttaatcag cagaccctcg ttggcgtaac ggaaccatga gaggtacgac aaccatttga 8640
ggtatactgg caccgagccc gagttcaaga agaaggcgtt tttccatagg ctccgccccc 8700
ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 8760
aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 8820
cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcaatgct 8880

```
cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg    8940
aacccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc    9000
cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga    9060
ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa    9120
ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta    9180
gctcttgatc cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc    9240
agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg    9300
acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga    9360
tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg    9420
agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct    9480
gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg    9540
agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc    9600
cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa    9660
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc    9720
cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt    9780
cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc    9840
ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt    9900
tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc    9960
catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt   10020
gtatgcggcg accgagttgc tcttgcccgg cgtcaacacg ggataatacc gcgccacata   10080
gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga   10140
tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag   10200
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa   10260
aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt   10320
attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga   10380
aaaataaaca aatagggttt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag   10440
aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc   10500
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca   10560
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg   10620
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc   10680
accatatgcg gtgtgaaata ccgcaccgaa tcgcgcggaa ctaacgacag tcgctccaag   10740
gtcgtcgaac aaaaggtgaa tgtgttgcgg agagcgggtg ggagacagcg aaagagcaac   10800
tacgaaacgt ggtgtggtgg aggtgaatta tgaagagggc gcgcgatttg aaaagtatgt   10860
atataaaaaa tatatcccgg tgttttatgt agcgataaac gagtttttga tgtaaggtat   10920
gcaggtgtgt aagtcttttg gttagaagac aaatccaaag tctacttgtg gggatgttcg   10980
aaggggaaat acttgtattc tataggtcat atcttgtttt tattggcaca aatataatta   11040
cattagcttt ttgaggggge aataaacagt aaacacgatg gtaataatgg taaaaaaaaa   11100
aacaagcagt tatttcggat atatgtcggc tactccttgc gtcgggcccg aagtcttaga   11160
gccagatatg cgagcacccg gaagctcacg atgagaatgg ccagac                  11206
```

<210> 2
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: 5'P end of vector, splice acceptor site

<400> 2
tttgcgagta cgcaaagctc tttctcttac ag 32

<210> 3
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: stop-start seq., Gal4 gene

<400> 3
gtcgaattga tgtgatggat ccaatgaagc t 31

<210> 4
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: mini-white gene

<400> 4
agacttcggg cccgacgcaa ggagtagaag 30

<210> 5
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: splice donor site, 3'P end of GT vector

<400> 5
gtaagtagcg gccgcacgta agggttaatg 30

<210> 6
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: anterior open cDNA exon 1

<400> 6
cagtacgtaa ccacagtgca cagtggagca aacag 35

<210> 7
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: anterior open cDNA exon 2

<400> 7
ataacaccgc ttaaaagcaa ttccagtggc ttccac 36